Reproductive Biology and Endocrinology

BioMed Central

Review

Open Access

Generating and manipulating transgenic animals using transposable elements

David A Largaespada*

Address: Department of Genetics, Cell Biology and Development, University of Minnesota Cancer Center, Minneapolis, MN 55455, USA

Email: David A Largaespada* - larga002@tc.umn.edu

* Corresponding author

Published: 07 November 2003

Reproductive Biology and Endocrinology 2003, **1**:80

Received: 15 July 2003

Accepted: 07 November 2003

This article is available from: http://www.rbej.com/content/1/1/80

Abstract

Transposable elements, or transposons, have played a significant role in the history of biological research. They have had a major influence on the structure of genomes during evolution, they can cause mutations, and their study led to the concept of so-called "selfish DNA". In addition, transposons have been manipulated as useful gene transfer vectors. While primarily restricted to use in invertebrates, prokaryotes, and plants, it is now clear that transposon technology and biology are just as relevant to the study of vertebrate species. Multiple transposons now have been shown to be active in vertebrates and they can be used for germline transgenesis, somatic cell transgenesis/gene therapy, and random germline insertional mutagenesis. The sophistication of these applications and the number of active elements are likely to increase over the next several years. This review covers the vertebrate-active retrotransposons and transposons that have been well studied and adapted for use as gene transfer agents. General considerations and predictions about the future utility of transposon technology are discussed.

Introduction

"One man's trash is another man's treasure." Anonymous

Many vertebrate genomes, including the human genome, are littered with trash, in the form of so-called "junk DNA". Only a small portion of most vertebrate genomes has an obvious biological role for the organism, such as protein encoding open reading frames, untranslated regions (UTR) of mRNA encoding DNA, promoter sequences, or ribosomal RNA encoding DNA. Instead, the vast majority of vertebrate genomic DNA has no obvious function and is indeed dominated by the presence of repetitive DNA elements that are themselves the remnants of transposable elements or the result of transposable element activity. Some of these elements have become inactive during the evolution of today's species, some are dependent upon enzymes encoded in trans by other transposable elements, and some remain active. These sequences are interesting for a variety of reasons. They can be used as markers of a species evolutionary history, they can be involved in the creation and destruction of genes, transposase genes have been "adopted" by some organisms to perform vital cellular functions (such as V-D-J recombination), and finally, transposons have been used as a source of vectors for transgenesis and mutagenesis in multiple species. It is this last feature of transposable elements, in particular their use in vertebrate species, that is the concern of this review. While the history of transposon study and vector development for use in bacteria, fungi, plant, and various invertebrate species is very long, it has been only very recently that transposable elements have demonstrated effectiveness as genetic tools in vertebrate species. We speculate that these elements will become important tools for vertebrate germline

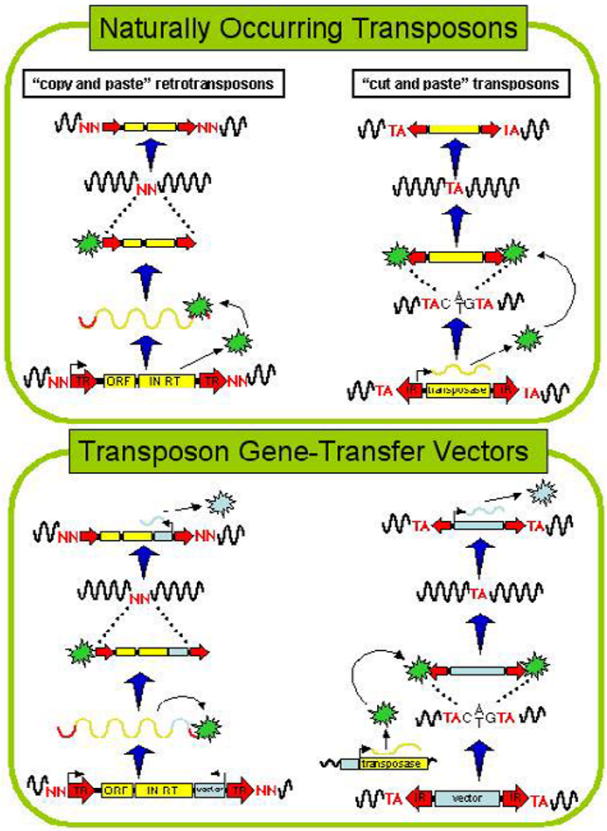

Figure 1
"Copy-and-paste" and "cut-and-paste" transposons have been adapted for use as gene transfer vectors. In the top half of the figure, transposition of naturally occurring transposons is depicted. In the lower half of the figure, the general methods used to adapt these transposons for use as gene transfer agents is shown. Direct terminal repeats (TR) flank some retrotransposons. Inverted terminal repeats (IR) flank cut and paste transposons. Retrotransposons, such as the L1 element, encode open reading frames (ORF) of unknown function as well as integrases (IN) and reverse transcriptases (RT). Both kinds of elements can be manipulated so that special vector sequences are inserted. In the case of retrotransposons, the vector sequences are inserted into the 3' untranslated region. In the case of the "cut and paste", DNA transposons, the vector sequences replace the transposase gene, which is expressed from a heterologous promoter in trans.

transgenesis, somatic and germline mutagenesis, and human gene therapy.

Transposable Elements

Many excellent reviews have been written on the subject of transposable elements and it is not the goal of this review to systematically cover that very large field, which really would require a textbook to do the subject justice [1–3]. It can be said, however, that transposons come in two general types (Figure 1). The "copy and paste" retrotransposons are mobilized by transcribing an RNA copy, that then becomes reverse transcribed and is integrated elsewhere in the genome. In contrast, the "cut and paste" transposable elements transpose by the direct excision from DNA and insertion elsewhere in the genome. Both types of elements have now been used in vertebrate cell lines and animals as gene transfer vectors. Retrotransposons come in many classes and include retroviruses, which will not be considered in this review. Retrotransposon elements require at a minimum, an internal promoter and coding sequences for an integrase and reverse transcriptase protein. Many of these elements encode proteins that act primarily in cis, on the RNA transcript from which they were translated. Thus, in biotechnology applications, foreign gene sequences must be added to the 3' untranslated region (UTR) of the retrotransposon vector (Figure 1). In contrast, the cut and paste transposases can act in trans. These elements have an internal promoter and encode a single protein "transposase" which binds to terminal repeat sequences on the ends of the element, causing it to be excised and inserted elsewhere. Because the cut and paste transposons can act in trans, it is feasible to supply the transposase by a variety of methods (DNA, RNA and possibly transferred protein) and insert any desired sequence in the transposon vector DNA itself (Figure 1). These sequences might include genes for germline or somatic cell transgenesis, sequences for insertional mutagenesis, or recognition by site-specific recombinases.

What follows is a review of recent success in developing vertebrate gene transfer vectors based on retrotransposons or DNA transposons. Currently these vector systems are at different levels of development and come from a variety of sources (Table 1). Some are elements taken from invertebrate species and adapted for use in vertebrates. Some are endogenous, naturally active vertebrate elements, while one was "reconstructed" from study of a large number of endogenous defective elements. Some of these vector systems have been shown to be effective for germline and somatic cell transgenesis *in vivo*, while others have so far only been shown to be active in cultured vertebrate cell lines. Finally, some have been introduced into transgenic mice and been mobilized from chromosomally resident positions. Thus, the transposons may be useful as general germline insertional mutagens. It is clear that multiple transposons, derived from various sources, might find utility for generating and manipulating transgenic animals. Because each vector system has advantages and disadvantages, depending on the specific application, multiple vector systems should be developed for use in the future.

Table 1: Transposable elements active in vertebrate species for use as gene transfer and insertional mutagenesis vectors.

Transposon	Family[a]	Activity in cultured cells[b]	Activity in vertebrate animals[c]	References
L1	LINE retrotransposon	Mouse (PCT) Human (PCT)	Mouse (TCRG)	[21,22]
Tol2	hAT	Mouse (EP) Human (EP)	Medaka (TCRG) Zebrafish (TCRG).	[27,28,32]
Tc1	Tc1/mariner	Human (PCT) Mouse (PCT)	-	[19,39,40]
Tc3	Tc1/mariner	Human (PCT)	Zebrafish (GT and TCRG)	[19,42]
Mariner (Himar1)	Tc1/mariner	Human (IPT)	-	[19,44]
Mariner (Mos1)	Tc1/mariner	Human (PCT)	Chicken (GT), zebrafish (GT)	[5,19,43]
Minos	Tc1/mariner	Human (PCT)	Mouse (TCRG) Mouse (TCRS)	[46–48]
Sleeping Beauty	Tc1/mariner	Mouse, hamster, human, monkey, dog, cow, sheep, quail, Xenopus, many fish (PCT)	Mouse (GT) Mouse lung, liver (SCT) Mouse (TCRG) Mouse (TCRS)	[4,19,30,51,53–56,59,60]

[a]Relevant superfamily of transposable element. [b]Activity in immortalized cell lines has been demonstrated using by demonstrating excision from an introduced plasmid (EP), interplasmid transposition (IPT), or full transposition from introduced plasmid (or viral vector) into chromosomes (PCT). [c]Activity in intact animals (various species) has been demonstrated by germline transgenesis (GT), somatic cell transgenesis (SCT), transposition of chromosomally resident transposon vectors in the germline (TCRG), and transposition of chromosomally resident transposon vectors in the soma (TCRS).

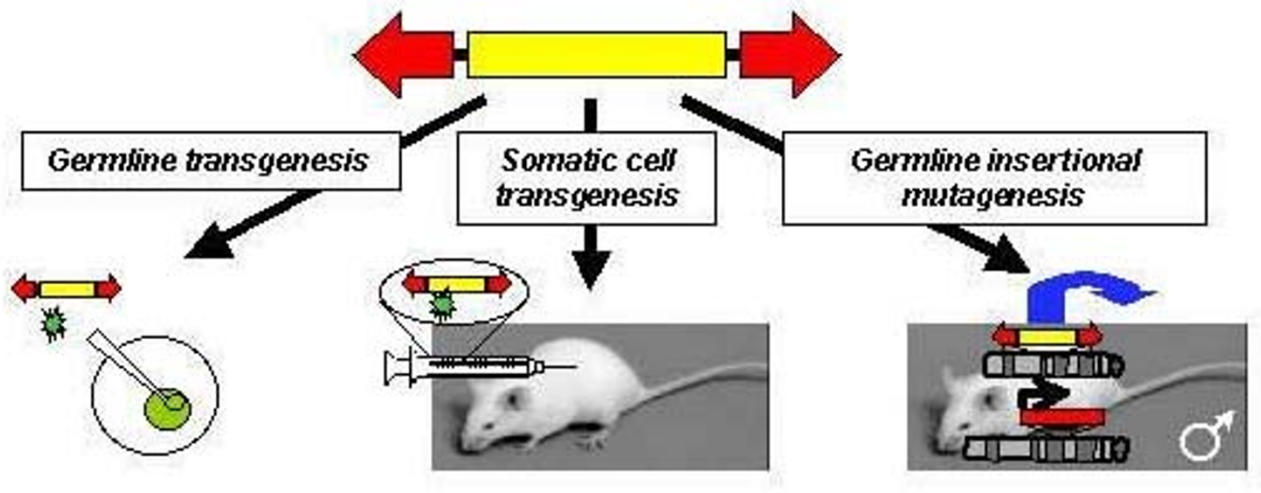

Figure 2
General uses for transposon vectors in the generation and manipulation of transgenic animals. Many uses can be imagined for transposon systems that are active in vertebrates. Three of the main uses are shown here.

General Considerations for Transposable Elements in Scientific and Biotechnology Applications

Transposons have the useful property of catalyzing the most important step in gene transfer applications, the insertion of foreign DNA into host chromosomes. A variety of uses for transposons in vertebrate science and biotechnology can be imagined. However, so far three general uses have been explored: germline transgenesis, somatic cell transgenesis/gene therapy, and random insertional mutagenesis (Figure 2).

First of all, germline transgenesis by transposition has been developed for certain of the transposon systems that are active in vertebrates. This has involved the co-delivery of *in vitro* transcribed mRNA encoding the transposase with transposon vector DNA into early embryos of the frog, *Xenopus tropicailis*, the zebrafish *Dana rario*, or the mouse (Paul Mead and Steve Ekker, personal communication) [4]. In the case of mariner-mediated transformation of the chicken germline, an active autonomous DNA element was injected [5]. It remains to be determined if

transgenesis in commercially important species such as the chicken, pig, goat or cow can be improved using transposons, but this area is certainly worth exploring given the success in diverse species so far. Given the complexities of harvesting fertilized embryos from large animals, thought also should be given to sperm transgenesis, since this may offer many practical benefits. Certain advantages can be imagined for achieving germline transgenesis by transposition, as compared to standard pronuclear injection of DNA. One advantage is the ability, in some cases, to generate offspring with multiple independent insertions that can then be segregated by breeding, thus increasing the number of total potentially useful transgenesis events. Standard pronuclear injection of linear DNA into embryos results in concatomerization of the DNA and random integration into the genome. However, these concatomer integration events are often associated with inversions, deletions or other large rearrangements of DNA at the integration site. Indeed, for this reason, up to 10% of all mouse germline transgenics have a distinct phenotype in the homozygous state due to insertional mutation/deletion of endogenous genes [6]. Although these events have allowed the identification of important endogenous genes, it is in general an undesired effect of germline transgenesis. The ability of transposons to deliver a specific fragment of DNA into a target site without alteration of endogenous sequences, other than the insertion of the transposon vector, could thus be considered an advantage. One disadvantage of transgenesis by transposition, could be the fact that multiple copies of a transgene cannot be integrated into one position by transgenesis, and so it may be difficult in some cases to achieve very high expression of transgenes by this method. Very low expression is a frequent problem for any transgenic project. Since transgene arrays are often subject to partial methylation and inactivation, as well as other mechanisms of silencing, often only a multicopy array can achieve expression levels rivaling the endogenous gene. That is, small transgene vectors typically express at a level that is only a fraction of the level of the native gene. Nevertheless, vertebrate germline transgenesis and expression has been achieved using transposons [4]. Important considerations for the use of transposons for germline transgenesis are the carrying capacity of the transposon vector, its ability to generate multiple, independent insertion events in one embryo, the tendency of introduced genes to be expressed within such vectors, and target site choice for integration. Obviously, vectors that can tolerate large DNA inserts have the advantage of allowing large cDNAs and regulatory sequences or multiple genes to be co-integrated at one position. As mentioned above, if multiple insertions per embryo are generated, then the total number of independent transgenesis events is increased, perhaps increasing the chance that one can be found which expresses at the desired level. Another important, but so far largely unexplored issue, is the effect of transposon vector sequences on transgene expression. It remains possible that inverted terminal repeat sequences from these vectors would be recognized by host cell machinery as repetitive sequence and thus genes in cis could be silenced by methylation. For this reason, and to avoid mobilization of endogenous elements, it may be best to use transposons in vertebrate species that have been isolated from very distantly related species. Finally, the tendency of a given transposon system to integrate vector DNA in transcribed versus non-transcribed chromatin may affect the overall transgene expression rate and tendency to mutate endogenous genes by insertion.

A second application of transposons is the introduction of transgenes directly into somatic cells of animals. Again, transposons offer the advantage of overcoming the most important barrier to gene transfer, that is, the stable integration of foreign DNA into the host cell chromosome. Much of this work done in this setting has been directed at eventual human gene therapy. Indeed, much of this work is very promising. It is worth considering other applications of somatic cell transgenesis however. In mouse cancer research, it sometimes would be desirable to avoid germline transgenesis and test the effect of transgenes directly in somatic cells. Not only might this allow many genes or gene versions to be tested for oncogenic ability, it also has the advantage of more faithfully mimicking human tumor development, in which genetically distinct clones are initiated and progress in an environment of normal tissue. Another application for somatic cell gene transfer could be in the area of biotechnology. It is usually assumed that animal bioreactors for the production of foreign proteins, for example in the milk of transgenic goats or cows, will be produced by germline transgenesis. However, it may be faster and more direct to introduce genes encoding these useful proteins, such as monoclonal antibodies, clotting factors, growth factors or other proteins, by introduction into appropriate somatic cells of large animals. Nevertheless, the largest body of work on transposons for somatic cell gene transfer has been done in mice, but for the eventual application in human gene therapy. The use of transposons for human gene therapy has some advantages over viral gene therapy and traditional non-viral or DNA-based gene therapy. Transposons offer the potential benefits of ease of pharmaceutical formulation and scale up since only DNA, or DNA and RNA encoding the transposase need to be produced. This approach may also produce vectors that are less immunogenic than are viral vectors making repeated administration more likely to be successful than with viral vectors. Again the issues of carrying capacity, target site preference for integration, and influence of transposon vector sequences on the expression of transgenes and

endogenous genes is critical for this application of transposon technology.

The third application of transposon technology is the insertional mutation of endogenous genes for genetic screens. This approach has been extensively pursued for genetic studies in invertebrates such as *Drosophila melanogaster* and in many plant species [7–10]. The advantage of such a system is that germline mutations caused by transposon insertion are molecularly tagged, greatly facilitating their molecular identification. This advantage is in stark contrast to germline mutagenesis using X-irradiation or chemicals, which often do not leave any easily identifiable trace of their activity or create such large chromosomal rearrangements or deletions as to affect multiple genes at one time. One of the best chemical mutagens for vertebrate germline mutagenesis is ethyl nitrosourea (ENU), an ethylating agents that generally causes single base pair mutations [11]. When used to treat fertile male mice for instance, ENU can mutate the average locus in one in every 750 to one in every 1500 gametes, making it an incredibly powerful mutagen [11]. Even so, each dominant or recessive mutation must be mapped to high resolution using meiotic recombination until a sufficiently small critical region of the genome is identified, usually much less than 1 cM in size. At this point, candidate genes in the region are laboriously screened for mutations by directly sequencing exons and splice junctions. Thus, gene identification remains a major bottleneck in ENU mutagenesis projects. Another advantage of transposon insertional mutagenesis is the ability to engineer special vectors that can express a reporter molecule in a context-dependent manner. Such vectors are often called "gene-traps" and can be designed to express a reporter gene if and only if insertion occurs in the right place and orientation. The three major types of gene-traps are: 1. enhancer traps, 2. promoter or 5' gene traps and 3. polyadenylation site or 3' gene traps. Enhancer traps contain a very weak promoter driving a reporter molecule, which if inserted near a gene can come under the influence of endogenous enhancer elements. The result is the temporal and/or tissue-specific expression of the reporter in transgenic animals. Enhancer trapping has been widely used in *Drosophila melanogaster* to create lines of flies expressing the GAL4 transcription factor in tissue specific patterns. In this way, a large number of so-called GAL4 driver lines can be bred with lines carrying other transgenes under the control of promoters containing upstream activating sequences (UAS) responsive to GAL4 [12]. Similar systems could be devised for use on vertebrate model organisms, perhaps based on GAL4/VP16 or tetracycline transactivator/VP16 fusions [13]. Promoter or 5' gene-traps created using transfected plasmid DNA or retroviral transduction have been widely pursued in cultured mouse embryonic stem (ES) cells to create large libraries of clones, each with one disrupted gene [14–17]. For this work, the vectors contain a splice acceptor followed by a gene conferring resistance to an antibiotic such as puromycin or G418. Thus, insertions into genes can be selected for in culture and the inserted gene is mutated since splicing into the plasmid or retroviral vector produces a truncated fusion transcript. Alternatively, 3' or polyadenylation site traps contain an internal promoter driving a reporter molecule or antibiotic resistance gene followed by a splice donor but lack any splice acceptor or polyadenylation sequence [17]. If the vector inserts into a gene-free region, then it's reporter protein will not be expressed, because the transcript is neither spliced nor polyadenylated, resulting in poor export from the nucleus and RNA instability. If the vector inserts within a gene and in the same orientation as that gene, then splicing to downstream exons results in the production of a stable fusion transcript that is spliced and polyadenylated, thus the reporter protein is expressed. In this way, ES clones can be selected that harbor insertions into genes, even if that gene is not expressed in ES cells. It should be noted that 3' traps are not thought to be highly mutagenic unless they contain an upstream 5' trap to efficiently truncate the endogenous gene [17]. Each of these three gene trapping vectors could be pursued for use in transposons for vertebrate germline mutagenesis. For this application, mobilization of chromosomally resident transposons in transgenic animals has the advantage of allowing new gene insertions to be obtained simply by breeding. Thus, true forward genetic screens could be performed in which phenotypes were identified first and then gene identifications made second. A critical issue for using any transposon for the purpose of germline insertional mutagenesis is the frequency of transposon mobilization. Too few insertions per gamete means that too few useful offspring are obtained, thus wasting animal resources and resulting in inefficiency. Too many insertions per gamete means that multiple, perhaps linked, genes could be mutated in the same animal. This will complicate identification of the insertion responsible for a given phenotype. Also critical for this application of transposon technology is a thorough understanding of the target site choice of the vector within the genome, with regard to the location of the donor site, transcribed versus non-transcribed regions, and within the genes into which it becomes inserted. For instance, some transposons, when mobilized from a chromosomal site, tend to insert within a local region near the donor site [18,19] (Carlson et al., Genetics, In Press). This effect will bias toward mutagenesis in a particular region of the genome. While this could be an advantage, local hopping would have an influence over a whole-genome screen.

What follows is a review of the transposable elements that have been shown to work in various applications

described above in vertebrate cell lines or animals. In the conclusion section, some predictions are made as to future development of these and other transposons for creating and manipulating transgenic animals.

L1 Elements

Among the vertebrate retrotransposable elements, the long interspersed repetitive elements (LINEs) have been extensively characterized and have been engineered for gene transfer and insertional mutagenesis [3]. LINEs are particularly abundant in the human genome with approximately 350,000 copies. Almost all of these elements are inactive due to mutations in the promoter or one or both of its two open reading frames. However, LINEs remain active in the human genome, and can in fact cause inherited genetic disorders [3]. It is estimated that up to 1% of human genetic mutations are due to LINE insertions. In contrast, mouse LINEs are much more active, causing perhaps 10% of mouse germline mutations [3,20]. The human L1 element promoter is active only in germline cells, but a heterologous promoter can be used to drive expression of the L1 transcript [21]. By cloning an L1 element, from an insertion into the factor VIII gene from a hemophilia A patient, Dr. Haig Kazazian's laboratory was able to isolate an active L1 element [21]. This vector was cloned and altered so that a reverse orientation, intron-interrupted, neomycin resistance gene (NEO) with a heterologous promoter was placed in the 3' UTR. After one round of transcription, splicing, reverse transcription, and integration, the NEO gene can be expressed and confers resistance to G418 selection. This cloned and altered L1 retrotransposon vector was shown to have activity in cultured human and mouse cells [21]. An altered version of this L1 retrotransposon was then constructed in which NEO was replaced by the gene for the green fluorescent protein (GFP) driven by the acrosin promoter [22]. In this way, retrotranspositon could be followed by the appearance of GFP positive cells. When transgenic mice were created with this L1/GFP element, an examination of sperm cells revealed a fairly high frequency of GFP+ cells. When these mice were bred to wild type females, approximately one in sixty offspring harbored new L1 vector insertions [22]. Potential advantages of this system include the fact that the entire genome potentially could be mutagenized using the L1 vector without bias to one chromosome or chromosome region, more insertions per genome might be obtained using recently identified more active L1 elements [23], and increases in transcription might provide increases in insertion rate. Current disadvantages include the relatively low number of insertions per gamete obtained so far, the fact that the 3' UTR of the L1 vector may tolerate only a limited amount of foreign sequence, the frequent 3' truncation of L1 vectors upon retrotransposition [21], and the fact that L1 insertion is often accompanied by deletion of sequences at the insertion site [24]. L1 vectors might be useful for introducing genes into somatic cells and have been cloned into adenoviral vectors, which can efficiently deliver L1 vectors to human cultured cell lines, resulting in retrotransposition into the genome [25].

hAT Elements: Tol2

In 1996, Dr. H. Hori's lab in Japan reported the isolation of a vertebrate transposon of the hobo-Activator-Tam3 (hAT) family of transposable elements inserted into the tryosinase gene of a strain of albino *Medaka* fish [26]. Subsequent research has shown that this element is indeed autonomous [27,28]. Tol2 is the first, and to date only, naturally occurring active cut-and-paste transposon isolated from vertebrates. However, an examination of endogenous Tc1/mariner-like zebrafish transposons suggests that some are active [29]. The Tol2 transposase gene is produced from a singly spliced transcript. The Tol2 transpoase gene can be produced from a heterologous promoter, while the transposon itself can be altered so that it contains a foreign gene cassette [27]. This two-part system can thus be used for a variety of gene transfer purposes. To date highly efficient germline transgenesis of the Zebrafish, *Dana rario*, has been achieved by co-injection of in vitro transcribed mRNA for the transposase with transposon vector DNA into the one cell embryo [27]. This method helps provide a system for efficient fish transgenesis, which is normally very inefficient by injection of plasmid DNA alone (Steve Ekker, personal communication). The Tol2 element may be active in mammalian cells, especially given the precedent for activity in human and mouse cells of a fish cut-and-paste transposon provided by study of *Sleeping Beauty* [4,30,31]. Indeed, Tol2 mediated excision, but not full transposition, has already been demonstrated in human and mouse cells [32]. It remains to be determined if the Tol2 transposase can be expressed as a transgene in the germline of transgenic animals and allow the mobilization of chromosomally resident transposon vectors. Potential advantages of Tol2 for germline mutagenesis include the fact that hAT elements have been shown to have the capacity to transpose very large vectors [33], and the fact that it is a naturally occurring element and so may be more active than reconstructed elements. Recent studies of the hobo element suggest that it has distinct hot spots for integration and that the percentage of the genome available for insertion is more restricted than for other transposable elements such as the P element [33,34].

Tc1/mariner Elements: Mariner (Himar1, Mos1), Tc1, Tc3, Minos, Sleeping Beauty

The Tc1/mariner family is a large family of widely distributed "cut-and-paste" transposable elements flanked by inverted terminal repeats, which in some cases themselves have embedded direct repeat sequences (for review see

[1]). These elements have been found in all vertebrate genomes examined, but are in all cases defective, containing frameshift mutations, stop codons and small or large deletions in the transposase gene. A lot of research has been invested into this family of transposons based upon investigations into the active proto-typical Tc1, from *Ceanorhabditis elegans*, and mariner elements, such as Mos1 from *Drosophila mauritania*. Members of this family contain a transposase with a DNA binding domain distantly related to the paired-box DNA binding domain. In addition, a catalytic DDE domain has been identified and critical amino acids in this domain have been discovered [1].

The mariner and Tc1 transposases have been purified, and remarkably, can be combined with a transposon DNA substrate, which will then undergo transposition *in vitro* without a requirement for any other proteins [35,36]. Two closely related versions of mariner have been well studied, Mos1 and Himar1, but are widespread in insects [37]. The Himar1 element is a consensus sequence based on several clones isolated from the horn fly, *Haematobia irritans*, and Mos1 was isolated from *Drosophila mauritania* [35]. The wide distribution of Tc1/mariner transposons, the ability of mariner to function in vitro in purified form and in distantly related species, suggested that such elements might require no co-factors for transposition in cells, or use very highly conserved host factors. For example, Mos1 has been introduced into the distantly related insect species, such as the yellow fever mosquito *Aedes aegypti* [38]. For these reasons, great enthusiasm exists for adapting Tc1/mariner-like transposons for use in vertebrate species. Indeed mariner, Tc1 and Tc3 show activity in cultured mouse and human cell lines [39,40]. A screen for hyperactive mutants in *E. coli* resulted in the identification of amino acid substitutions that improve the activity of Himar1 in cultured cells [41]. The Mos1 mariner transposon has been used to achieve germline transgenesis in the chicken [5]. Both mariner and Tc3 has been used to achieve germline transgenesis in zebrafish [42,43]. The purified Himar1 transposase efficiently catalyzes interplasmid transposition [35]. The same kind of interplasmid transposition assay was used in a human 293T embryonic kidney cell line, showing that Himar1 transposition can occur in mammalian cells [44]. While Himar1 is primarily used for insertional mutagenesis in bacteria and mycobacteria, its potential for vertebrate gene transfer and insertional mutagenesis remains unclear. It remains to be determined if the active mariner elements, Tc1, or Tc3 can be used to mobilize chromosomally resident transposon vectors in vertebrates.

The Minos element, a Tc1/mariner-like transposon from *Drosophila hydei*, has been shown to be active in cultured human cell lines [45,46]. A gene-trap Minos transposon has been constructed and a library of insertions in HeLa cells were obtained [46]. The Minos transpoase gene has been expressed in transgenic mice from B cell lineage specific and male sperm cell specific promoters [47,48]. When these mice were crossed with transgenic mice carrying Minos transposon vectors, transposition occurred in B cells and in the male germline respectively. Germline insertions occurred in rough one in every ten offspring. Several germline insertions were cloned and each was found to be on a different chromosome [48]. These data suggest that Minos might be useful for mouse germline mutagenesis and that no distinct preference for local transposition occurs using this transposon system. Thus, genome-wide insertional mutagenesis with Minos is a real possibility in the mouse and possibly other species as well.

The *Sleeping Beauty* transposon (SB) is a synthetic Tc1/mariner family transposon derived from defective elements cloned from various Salmonid fish genomes [30]. The development of SB is particularly significant because it is the first vertebrate transposable element reconstructed from defective endogenous elements. This was accomplished by the stepwise repair of the open reading frame, nuclear localization signal, DNA binding domain, and catalytic activity. Because this process required ten major steps, the SB transposase was designated SB10. SB has been extensively studied. A model of the SB transposition reaction has been proposed based on studies of its inverted terminal repeat structures [49]. As part of this study an improved inverted terminal repeat sequence, designated pT2, was discovered [49]. SB related transposons have complicated inverted terminal repeats, each with two embedded direct repeats called the inner direct repeats and the outer direct repeats. The direct repeats, which are ~25 base pairs long, are the sites of transposase binding [30]. Neither the inverted repeats nor the direct repeats are perfect, and clear differences exist between the right and left inverted repeats and between the inner and outer direct repeats. The rules governing the transposition are not completely understood, but higher binding activity does not translate into increased transposition [49]. In addition, continued examination of the SB10 transposase sequence led to the identification of additional amino acid substitutions that confer increased activity in gene transfer into transfected HeLa cells [50]. SB transposons have been used to achieve germline transgenesis in the mouse [4]. In these experiments, one-cell mouse embryo pronuclei were co-injected with transposon vector DNA and in vitro transcribed mRNA encoding the SB10 transposase. The overall transgenesis rate was increased 1.5 fold over background without transposase. The increase in transgenesis rate was entirely due to offspring with multiple, independent transposon insertions. An average of three insertions per animal was obtained in some experi-

ments and transposon transgenes could be expressed. However, the transgenes within transposon vectors seem to be subject to position effect variagation, just as are standard mouse transgenes. These data showed that mammalian germline transgenesis by transposition is possible and suggest that the SB system might find utility in other mammalian species in which transgenesis is difficult. Another application of SB has been somatic cell transgenesis. SB can be used for stable, long-term gene transfer and expression into the liver of adult mice [51]. In some experiments, the SB transposon vector and transposase transgene have been delivered by so-called "hydrodynamic therapy". In this approach, 10% of the weight of the mouse of DNA-containing Ringer's solution is injected via the tail vein into mice in less than ten seconds. For the average mouse this is roughly 2–2.5 milliliters. The resultant transient increase in venous pressure within the liver is thought to result in extravasation of the plasmid DNA, which is efficiently taken up by ~25% of hepatocytes [52]. The Factor IX gene was cloned into an SB transposon vector and delivered to hemophilia B knockout mice using this technique by Dr. Mark Kay's laboratory [51]. Long-term expression required co-delivery of the Factor IX transposon and an active SB10 transposase gene on another plasmid. Similar success has been achieved for long-term gene transfer of the FAH gene into knockout mouse liver [53], and the *LAMB3* gene into cultured human skin cells from patients with junctional epidermolysis bullosa syndrome to correct this disorder in xenografted nude mice [54]. Dr. Mark Kay's group also has developed "binary" gene therapy vectors in which both the SB transposon vector and SB10 transposase gene were delivered to hepatocytes using adenoviral vectors, which by themselves only result in transient gene transfer and expression [31]. Interestingly, efficient gene transposition required recombinase-mediated excision and circularization of the transposon vector from the linear adenoviral DNA, suggesting that circularized transposon vector DNA is more efficiently transposed, at least in this environment [31]. Dr. Scott McIvor's laboratory has developed methods for long term gene transfer and expression into adult mouse lung using SB vector and SB10 transposase DNA complexed with polyethyleneimine (PEI), a polycationic, branched molecule (Beleur et al., Molecular Therapy, In Press). Finally, in the area of germline insertional mutagenesis, we have demonstrated that SB transposons present in chromosomes of SB transposase transgenic mice will efficiently undergo transposition in germline cells, such that offspring from these mice are obtained with new transposon insertions [55]. Similar results were published in 2001 by two other labs using SB, one in Japan and one in the Netherlands [19,56]. Three papers all report essentially similar results, with differences in the average number of new transposon insertions obtained. The ubiquitously expressed CAGGS [55,56], or the male germline specific protamine 1 promoters [19] were used to drive expression of the SB transposase. Single copy [19] or multicopy transposon vectors were used, all with different internal sequences in the transposon itself. An average of 0.2 [19], 1 [56], or 2.0 [55] new transposon insertions per offspring were obtained. These data clearly establish that the SB system can be used to achieve high efficiency transposition in the male or female germline. When we bred animals with new transposon insertions they were present in roughly half of the offspring of these animals and could segregate independently as if they had transposed to multiple chromosomes. Moreover, if the animal bred was also transgenic for the SB10 transposase, then a number of new transposon insertions were detected. It is important to note that offspring are generated with as many as 11 new transposon insertions, when the transposon concatomer is passed through the germline twice in the presence of the transposase transgene (unpublished data). To be useful as insertional mutagens, SB transposon vectors should be capable of inserting into genes. Indeed, transposon vector insertion into genes has been observed in primary mouse liver cells [51] and in cultured cell lines. Using inverse PCR or a linker-mediated PCR technique [55] we have cloned and sequenced 44 germline transposon insertions and analyzed them and mice carrying these insertions (Carlson et al., Genetics, In Press). All the insertions cloned are flanked by TA dinucleotides as expected for Tc1/mariner family transposition. Analysis of the transposon insertions showed that the adjacent plasmid sequence from the concatomer had been replaced by mouse genomic sequence as expected if true transposition had occurred. The distribution and sequence content flanking these cloned insertion sites was compared to 44 mock insertion sites randomly selected from the genome. We found that germline SB transposon sites are AT-rich and the sequence ANNTANNT is favored compared to other TA dinucleotides. Local transposition occurs with insertions linked to the donor site roughly 40% of the time. The size of this local hopping interval is roughly 3–5 cM or 10–12 Mbp. We find roughly 30% of the transposon insertions are in transcription units as determined using the Celera database, similar to the percentage of random TA dinucleotides. We also determined that transposons inserted within a gene, in the same orientation as the gene, are subject to splicing from upstream exons of the endogenous gene. Significantly, we now know how often transposon insertions occur within genes (~25–30% of the time), how often a transposon insertion occurs locally near the donor site (~50% of the time), and how big the local region is for SB-mediated transposition (~3–5 cM or 10–15 Mbp). The results are all consistent with the use of SB for forward genetic screens in local intervals of the mouse genome. Thus, assuming 2 new transposon insertions per gamete (as we have already achieved) we can

expect to achieve a 1X coverage of a 10 Mbp region of the genome (with one insertion every 10 kb) in as few as 1000 mice. We have begun to characterize two embryonic lethal mutations caused by endogenous splicing disruption in mice carrying intron-inserted SB gene-trap transposons. It is clear from these analyses that SB and probably other cut-and-paste transposons have potential utility as random germline mutagens for forward genetic screens.

Conclusions and Future Directions

It is very likely that transposon technology will have an impact in one or more of the three applications described in this review over the next several years. This progress may involve use of one or more of the transposon systems described above or may involve new transposon systems. The precedent set by work on SB shows that evolutionarily defunct transposons can be "resurrected" using reverse evolutionary principals. This means that many other potentially useful transposons could be derived from a purely informatics based approach using sequenced genomes. A rich source of new transposon systems could be generated in this way, some of which may have attributes more suited to one or another application. Alternately, it is likely that many species contain active elements, such as Tol2 and L1, since we see the evidence of their past activity in all genomes examined. Again, the search for these active elements will be a useful outcome of ongoing genome projects. Mammalian or avian germline transgenesis by transposition could have an impact on several agriculturally relevant species. Transgenesis for most of these species is currently very difficult or impossible. Given the technical challenges of harvesting, manipulating, and injecting early embryos from some of these species, it is worth considering sperm transgenesis by transposition. This process has been achieved in mice via micro-injection [57], and might be made more efficient using transposon vectors. While the risk of insertional mutagenesis, specifically activation of endogenous proto-oncogenes and cancer, are present with transposons used in gene therapy [58], it remains to be determined if they are low of enough compared to the benefits that transposon-based gene therapy may bring. However, it is clear from work using SB that the concept is sound from a technical standpoint. Finally, transposon-mediated insertional mutagenesis of the mouse germline is clearly possible with SB, Minos and probably Tol2. Transposon-based germline mutagenesis might also be considered for other species, particularly those for which ES cells cannot be easily obtained. In the mouse, however, the best use of each system probably depends upon the type of screen desired. The maximal number of insertions per gamete that can be reliably obtained must be determined. The ideal sequences for gene-trapping must be identified. Applications such as local saturation mutagenesis or chromosome engineering by placement of LoxP sites within transposon vectors are certainly possible. It is hoped that such approaches can led to significant contributions to functional annotation of the mammalian genome in the future.

Acknowledgments

The author thanks Drs. Perry Hackett, Steve Ekker, Scott McIvor, and other members of the University of Minnesota Arnold and Mabel Beckman Center for Transposon Research for helpful suggestions and technical review. This work was partially supported by The Arnold and Mabel Beckman Foundation, the University of Minnesota Academic Health Center and Graduate School, and NIH grant R01 DA14764.

References

1. Plasterk RH, Izsvak Z and Ivics Z: **Resident aliens: the Tc1/mariner superfamily of transposable elements.** *Trends Genet* 1999, **15:**326-332.
2. Spradling AC, Stern DM, Kiss I, Roote J, Laverty T and Rubin GM: **Gene disruptions using P transposable elements: an integral component of the Drosophila genome project.** *Proc Natl Acad Sci U S A* 1995, **92:**10824-10830.
3. Ostertag EM and Kazazian HH Jr: **Biology of mammalian L1 retrotransposons.** *Annu Rev Ge* 2001, **35:**501-538.
4. Dupuy AJ, Clark K, Carlson CM, Fritz S, Davidson AE, Markley KM, Finley K, Fletcher CF, Ekker SC, Hackett PB, Horn S and Largaespada DA: **Mammalian germ-line transgenesis by transposition.** *Proc Natl Acad Sci U S A* 2002, **99:**4495-4499.
5. Sherman A, Dawson A, Mather C, Gilhooley H, Li Y, Mitchell R, Finnegan D and Sang H: **Transposition of the Drosophila element mariner into the chicken germ line.** *Nat Biotechnol* 1998, **16:**1050-1053.
6. Babinet C, Morello D and Renard JP: **Transgenic mice.** *Genome* 1989, **31:**938-949.
7. Spradling AC and Rubin GM: **Transposition of cloned P elements into Drosophila germ line chromosomes.** *Science* 1982, **218:**341-347.
8. Spradling AC, Stern D, Beaton A, Rhem EJ, Laverty T, Mozden N, Misra S and Rubin GM: **The Berkeley Drosophila Genome Project gene disruption project: Single P-element insertions mutating 25% of vital Drosophila genes.** *Genetics* 1999, **153:**135-177.
9. Greco R, Ouwerkerk PB, Sallaud C, Kohli A, Colombo L, Puigdomenech P, Guiderdoni E, Christou P, Hoge JH and Pereira A: **Transposon Insertional Mutagenesis in Rice.** *Plant Physiol* 2001, **125:**1175-1177.
10. Walbot V: **Saturation mutagenesis using maize transposons.** *Curr Opin Plant Biol* 2000, **3:**103-107.
11. Justice MJ, Noveroske JK, Weber JS, Zheng B and Bradley A: **Mouse ENU mutagenesis.** *Hum Mol Genet* 1999, **8:**1955-1963.
12. Duffy JB: **GAL4 system in Drosophila: a fly geneticist's Swiss army knife.** *Genesis* 2002, **34:**1-15.
13. Furth PA, St Onge L, Boger H, Gruss P, Gossen M, Kistner A, Bujard H and Hennighausen L: **Temporal control of gene expression in transgenic mice by a tetracycline-responsive promoter.** *Proc Natl Acad Sci U S A* 1994, **91:**9302-9306.
14. Hicks GG, Shi EG, Li XM, Li CH, Pawlak M and Ruley HE: **Functional genomics in mice by tagged sequence mutagenesis.** *Nat Genet* 1997, **16:**338-344.
15. Skarnes WC, Moss JE, Hurtley SM and Beddington RS: **Capturing genes encoding membrane and secreted proteins important for mouse development.** *Proc Natl Acad Sci U S A* 1995, **92:**6592-6596.
16. Townley DJ, Avery BJ, Rosen B and Skarnes WC: **Rapid sequence analysis of gene trap integrations to generate a resource of insertional mutations in mice.** *Genome Res* 1997, **7:**293-298.
17. Zambrowicz BP, Friedrich GA, Buxton EC, Lilleberg SL, Person C and Sands AT: **Disruption and sequence identification of 2,000 genes in mouse embryonic stem cells.** *Nature* 1998, **392:**608-611.
18. Tower J, Karpen GH, Craig N and Spradling AC: **Preferential transposition of Drosophila P elements to nearby chromosomal sites.** *Genetics* 1993, **133:**347-359.

19. Fischer SE, Wienholds E and Plasterk RH: **Regulated transposition of a fish transposon in the mouse germ line.** *Proc Natl Acad Sci U S A* 2001, **98:**6759-6764.
20. Naas TP, DeBerardinis RJ, Moran JV, Ostertag EM, Kingsmore SF, Seldin MF, Hayashizaki Y, Martin SL and Kazazian HH: **An actively retrotransposing, novel subfamily of mouse L1 elements.** *Embo J* 1998, **17:**590-597.
21. Moran JV, Holmes SE, Naas TP, DeBerardinis RJ, Boeke JD and Kazazian HH: **High frequency retrotransposition in cultured mammalian cells.** *Cell* 1996, **87:**917-927.
22. Ostertag EM, DeBerardinis RJ, Goodier JL, Zhang Y, Yang N, Gerton GL and Kazazian HH Jr: **A mouse model of human L1 retrotransposition.** *Nat Genet* 2002, **32:**655-660.
23. Brouha B, Schustak J, Badge RM, Lutz-Prigge S, Farley AH, Moran JV and Kazazian HH Jr: **Hot L1s account for the bulk of retrotransposition in the human population.** *Proc Natl Acad Sci U S A* 2003, **100:**5280-5285.
24. Gilbert N, Lutz-Prigge S and Moran JV: **Genomic deletions created upon LINE-1 retrotransposition.** *Cell* 2002, **110:**315-325.
25. Soifer H, Higo C, Kazazian HH Jr, Moran JV, Mitani K and Kasahara N: **Stable integration of transgenes delivered by a retrotransposon-adenovirus hybrid vector.** *Hum Gene Ther* 2001, **12:**1417-1428.
26. Koga A, Suzuki M, Inagaki H, Bessho Y and Hori H: **Transposable element in fish.** *Nature* 1996, **383:**30.
27. Kawakami K, Shima A and Kawakami N: **Identification of a functional transposase of the Tol2 element, an Ac-like element from the Japanese medaka fish, and its transposition in the zebrafish germ lineage.** *Proc Natl Acad Sci U S A* 2000, **97:**11403-11408.
28. Koga A and Hori H: **Detection of de novo insertion of the medaka fish transposable element Tol2.** *Genetics* 2000, **156:**1243-1247.
29. Lam WL, Lee TS and Gilbert W: **Active transposition in zebrafish.** *Proc Natl Acad Sci U S A* 1996, **93:**10870-10875.
30. Ivics Z, Hackett PB, Plasterk RH and Izsvak Z: **Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells.** *Cell* 1997, **91:**501-510.
31. Yant SR, Ehrhardt A, Mikkelsen JG, Meuse L, Pham T and Kay MA: **Transposition from a gutless adeno-transposon vector stabilizes transgene expression in vivo.** *Nat Biotechnol* 2002, **16:**16.
32. Koga A, Iida A, Kamiya M, Hayashi R, Hori H, Ishikawa Y and Tachibana A: **The medaka fish Tol2 transposable element can undergo excision in human and mouse cells.** *J Hum Genet* 2003, **48:**231-235.
33. Ladeveze V, Aulard S, Chaminade N, Biemont C, Periquet G and Lemeunier F: **Dynamics of the hobo transposable element in transgenic lines of Drosophila melanogaster.** *Genet Res* 2001, **77:**135-142.
34. Guimond N, Bideshi DK, Pinkerton AC, Atkinson PW and O'Brochta DA: **Patterns of Hermes transposition in Drosophila melanogaster.** *Mol Genet Genomics* 2003, **268:**779-790.
35. Lampe DJ, Churchill ME and Robertson HM: **A purified mariner transposase is sufficient to mediate transposition in vitro.** *Embo J* 1996, **15:**5470-5479.
36. Vos JC, De Baere I and Plasterk RH: **Transposase is the only nematode protein required for in vitro transposition of Tc1.** *Genes Dev* 1996, **10:**755-761.
37. Robertson HM: **The mariner transposable element is widespread in insects.** *Nature* 1993, **362:**241-245.
38. Coates CJ, Jasinskiene N, Miyashiro L and James AA: **Mariner transposition and transformation of the yellow fever mosquito, Aedes aegypti.** *Proc Natl Acad Sci U S A* 1998, **95:**3748-3751.
39. Schouten GJ, van Luenen HG, Verra NC, Valerio D and Plasterk RH: **Transposon Tc1 of the nematode Caenorhabditis elegans jumps in human cells.** *Nucleic Acids Res* 1998, **26:**3013-3017.
40. Li ZH, Liu DP, Wang J, Guo ZC, Yin WX and Liang CC: **Inversion and transposition of Tc1 transposon of C. elegans in mammalian cells.** *Somat Cell Mol Genet* 1998, **24:**363-369.
41. Lampe DJ, Akerley BJ, Rubin EJ, Mekalanos JJ and Robertson HM: **Hyperactive transposase mutants of the Himar1 mariner transposon.** *Proc Natl Acad Sci U S A* 1999, **96:**11428-11433.
42. Raz E, van Luenen HG, Schaerringer B, Plasterk RH and Driever W: **Transposition of the nematode Caenorhabditis elegans Tc3 element in the zebrafish Danio rerio.** *Curr Biol* 1998, **8:**82-88.
43. Fadool JM, Hartl DL and Dowling JE: **Transposition of the mariner element from Drosophila mauritiana in zebrafish.** *Proc Natl Acad Sci U S A* 1998, **95:**5182-5186.
44. Zhang L, Sankar U, Lampe DJ, Robertson HM and Graham FL: **The Himar1 mariner transposase cloned in a recombinant adenovirus vector is functional in mammalian cells.** *Nucleic Acids Res* 1998, **26:**3687-3693.
45. Franz G, Loukeris TG, Dialektaki G, Thompson CR and Savakis C: **Mobile Minos elements from Drosophila hydei encode a two-exon transposase with similarity to the paired DNA-binding domain.** *Proc Natl Acad Sci U S A* 1994, **91:**4746-4750.
46. Klinakis AG, Zagoraiou L, Vassilatis DK and Savakis C: **Genome-wide insertional mutagenesis in human cells by the Drosophila mobile element Minos.** *EMBO Rep* 2000, **1:**416-421.
47. Zagoraiou L, Drabek D, Alexaki S, Guy JA, Klinakis AG, Langeveld A, Skavdis G, Mamalaki C, Grosveld F and Savakis C: **In vivo transposition of Minos, a Drosophila mobile element, in mammalian tissues.** *Proc Natl Acad Sci U S A* 2001, **98:**11474-11478.
48. Drabek D, Zagoraiou L, deWit T, Langeveld A, Roumpaki C, Mamalaki C, Savakis C and Grosveld F: **Transposition of the drosophila hydei minos transposon in the mouse germ line.** *Genomics* 2003, **81:**108-111.
49. Cui Z, Geurts AM, Liu G, Kaufman CD and Hackett PB: **Structure-function analysis of the inverted terminal repeats of the sleeping beauty transposon.** *J Mol Biol* 2002, **318:**1221-1235.
50. Geurts AM, Yang Y, Clark KJ, Liu G, Cui Z, Dupuy AJ, Bell JB, Largaespada DA and Hackett PB: **Gene transfer into genomes of human cells by the sleeping beauty transposon system.** *Mol Ther* 2003, **8:**108-117.
51. Yant SR, Meuse L, Chiu W, Ivics Z, Izsvak Z and Kay MA: **Somatic integration and long-term transgene expression in normal and haemophilic mice using a DNA transposon system.** *Nat Genet* 2000, **25:**35-41.
52. Stoll SM, Sclimenti CR, Baba EJ, Meuse L, Kay MA and Calos MP: **Epstein-Barr virus/human vector provides high-level, long-term expression of alpha1-antitrypsin in mice.** *Mol Ther* 2001, **4:**122-129.
53. Montini E, Held PK, Noll M, Morcinek N, Al-Dhalimy M, Finegold M, Yant SR, Kay MA and Grompe M: **In Vivo Correction of Murine Tyrosinemia Type I by DNA-Mediated Transposition.** *Mol Ther* 2002, **6:**759-769.
54. Ortiz-Urda S, Lin Q, Yant SR, Keene D, Kay MA and Khavari PA: **Sustainable correction of junctional epidermolysis bullosa via transposon-mediated nonviral gene transfer.** *Gene Ther* 2003, **10:**1099-1104.
55. Dupuy AJ, Fritz S and Largaespada DA: **Transposition and gene disruption in the male germline of the mouse.** *Genesis* 2001, **30:**82-88.
56. Horie K, Kuroiwa A, Ikawa M, Okabe M, Kondoh G, Matsuda Y and Takeda J: **Efficient chromosomal transposition of a Tc1/mariner-like transposon Sleeping Beauty in mice.** *Proc Natl Acad Sci U S A* 2001, **98:**9191-9196.
57. Perry AC, Wakayama T, Kishikawa H, Kasai T, Okabe M, Toyoda Y and Yanagimachi R: **Mammalian transgenesis by intracytoplasmic sperm injection.** *Science* 1999, **284:**1180-1183.
58. Verma IM: **Success and setback: another adverse event.** *Mol Ther* 2002, **6:**565-566.
59. Luo G, Ivics Z, Izsvak Z and Bradley A: **Chromosomal transposition of a Tc1/mariner-like element in mouse embryonic stem cells.** *Proc Natl Acad Sci U S A* 1998, **95:**10769-10773.
60. Izsvak Z, Ivics Z and Plasterk RH: **Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates.** *J Mol Biol* 2000, **302:**93-102.